



Guest Editorial: Computational Intelligence for Big Data Analytics

A. Fernandez¹ · I. Triguero² · M. Galar³ · F. Herrera¹

Received: 27 March 2019 / Accepted: 10 May 2019 / Published online: 25 May 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Nowadays, we are witnessing exponential growth in the generation of raw data. This new scenario comes with several challenges and opportunities when managing, analyzing, synthesizing, visualizing and, in summary, discovering underlying knowledge within this massive amount of information. The set of technologies and methodologies to address this task has been defined as Big Data Analytics [4, 8].

One of the key objectives of Big Data applications is to provide human-centered smart systems. This was made possible thanks to the rapid development of emerging areas such as Cognitive Computing [1]. Specifically, it comprises a fusion of several technologies, including networking, analytics, and cloud computing, aiming to obtain an automatic computational system capable of exhibiting some degree of brain-like cognitive intelligence [5, 6].

The benefits derived from the use of Big Data Analytics come along with the need for designing scalable solutions [7]. Among different learning paradigms, in this special

issue, we focus on Computational Intelligence (CI) models [2], especially those based on fuzzy sets and fuzzy logic, evolutionary computation, and their hybridization in terms of Evolutionary Fuzzy Systems (EFSs) [3].

This guest special issue encompasses seven papers devoted to recent developments in the field of CI solutions for Big Data. It was originated from a series of invited special sessions organized by the guest editors at the 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2017) that was held in Naples, Italy, July 9–12, 2017.

These seven original contributions cover several topics such as classification, regression, association rule mining and emerging patterns. Additionally, some real-world applications show the true benefits of the usage of CI solutions in the context of Big Data. All the research studies have been thoroughly revised and extended to become the papers currently presented in this issue. These proposals are briefly described as follows.

The paper “A grammar-guided genetic programming algorithm for associative classification in Big Data” by F. Padillo, J.M. Luna, and S. Ventura introduces a novel approach in the context of associative classification. Instead of considering the classical four stages of the procedure, including data discretization plus pattern mining as the most computationally expensive ones, the authors develop a two phase model via a grammar-guided genetic programming framework. It works by first mining reliable association rules, and then obtaining the classifier by ranking and combining the previously mined rules. Another interesting point of this approach is the implementation that considers different architectures such as multi-thread, Apache Spark, and Apache Flink.

In the paper “An incremental approach to address Big Data classification problems using fuzzy rules” by A. Gonzalez, R. Perez, and R. Romero-Zaliz, the authors present an alternative methodology to cope with large scale problems based on a simplified sequential iterative learning, instead of a general MapReduce scheme. By means of an

✉ A. Fernandez
alberto@decsai.ugr.es

I. Triguero
Isaac.Triguero@nottingham.ac.uk

M. Galar
mikel.galar@unavarra.es

F. Herrera
herrera@decsai.ugr.es

¹ DaSCI Andalusian Institute of Data Science and Computational Intelligence, University of Granada, Granada, Spain

² The Automated Scheduling Optimisation and Planning Research Group, School of Computer Science, University of Nottingham, Nottingham, UK

³ Institute of Smart Cities, Public University of Navarre, Pamplona, Spain

evolutionary process, the algorithm is capable of learning few high quality fuzzy rules, avoiding the problem of locality from the traditional local distributed models.

The paper “Optimizing partition granularity, membership function parameters, and rule bases of fuzzy classifiers for Big Data by a multi-objective evolutionary approach” by M. Barsacchi et al. presents an EFS with special focus on the interpretability of the system. More specifically, a Spark-based implementation of a fuzzy decision tree as learning stage, followed by an evolutionary tuning of the granularity and rule selection is presented. The final model is shown to be as accurate as the original EFS with a much lower degree of complexity.

The paper “Designing scalable linguistic fuzzy rule-based regression systems with evolutionary adaptive defuzzification” by S. Lopez et al. is focused on the tasks of regression and control. In particular, it proposes a scalable evolutionary adaptive defuzzification for increasing the quality of the final fuzzy model. This solution makes use of the MapReduce scheme within Spark.

In the paper “A Big Data approach for the extraction of fuzzy emerging patterns,” A.M. Garcia-Vico et al. proposed a multi-objective evolutionary algorithm for emerging pattern mining in Big Data environments. This recent Machine Learning task consists of obtaining a descriptive model aimed at finding discriminant characteristics in the data. In this research, authors show how a MapReduce-based model is capable of efficiently tackling very large datasets with promising results.

The paper “Feature selection and evolutionary rule learning for Big Data in smart building energy management”, by P. Rodriguez, M. Mucientes, and A. Bugarin is based on a real-case application. The authors focused on achieving an efficient energy management for smart buildings, as it may provide significant cost reductions and less environmental impact. In their research, a scalable EFS for understanding and predicting the future consumption of the building is presented. The proposed solution is capable of providing an accurate and interpretable model integrating feature selection and a Big Data regression learning scheme.

The paper entitled “PAS3-HSID: a dynamic bio-inspired approach for real-time hot spot identification in data streams” by R. Tickle et al. addresses some of the challenges derived from detecting hot spots in massive amounts of streaming data. Hot spots identification problems are present in a wide variety of areas such as health care, energy or transportation, and consist of determining a particular area with a high likelihood of occurrence of a certain event. In this article, authors proposed a dynamic bio-

inspired approach to handle Big Data streams using Apache Spark Streaming, showing a great potential in a real-world transportation problem.

The guest editors hope that the seven contributions of this special issue will trigger further advances of CI in the field of Big Data and Data Science. Finally, the guest editors wish to deeply thank all the authors for their contributions and the reviewers for their outstanding cooperation. We sincerely thank A. Hussain, present Editor-in-Chief of the Cognitive Computation journal, for providing us with the opportunity to edit this special issue.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval This article does not contain any studies with human participants performed by any of the authors.

References

1. Chen M, Herrera F, Hwang K. Cognitive computing: architecture, technologies and intelligent applications. *IEEE Access*. 2018;6:19774–83.
2. Eberhart RC, Shi Y. *Computational intelligence*, 1st edn. San Mateo: Morgan Kaufmann; 2007.
3. Fernandez A, del Jesus MJ, Cordon O, Marcelloni F, Herrera F. Evolutionary fuzzy systems for explainable artificial intelligence: why, when, what for, and where to *IEEE Comput Intell Mag*. 2019;14(1):69–81.
4. Fernández A, Río S, López V, Bawakid A, del Jesus MJ, Benítez JM, Herrera F. Big data with cloud computing: an insight on the computing environment, MapReduce and programming framework. *WIREs Data Min. Knowl. Discovery*. 2014;4(5):380–409.
5. Howard N, Hussain A. The fundamental code unit of the brain: Towards a new model for cognitive geometry. *Cogn. Comput*. 2018;10(3):426–36.
6. Mahmud M, Kaiser M, Rahman M, Rahman M, Shabut A, Al-Mamun S, Hussain A. A brain-inspired trust management model to assure security in a cloud based IoT framework for neuroscience applications. *Cogn. Comput*. 2018;10(5):864–73.
7. Ramírez-Gallego S, Fernández A, García S, Chen M, Herrera F. Big data: Tutorial and guidelines on information and process fusion for analytics algorithms with MapReduce. *Inf Fusion*. 2018;42:51–61.
8. Rao TR, Mitra P, Bhatt R, Goswami A. The big data system, components, tools, and technologies: a survey. *Knowl Inf Syst*. 2019;1–81. <https://doi.org/10.1007/s10115-018-1248-0>.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.